

NIH Genomic Data Sharing:

Managing Variables & Risks in an Evolving Landscape

Laura Lyman Rodriguez, Ph.D.

Office of Policy, Communications and Education

National Human Genome Research Institute

Presidential Commission for the Study of Bioethical Issues

August 1, 2012

NATIONAL
HUMAN GENOME
RESEARCH INSTITUTE



The NIH GWAS Data Sharing Policy

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS)

AGENCY: National Institutes of Health, HHS.

ACTION: Notice.

Background

The NIH is interested in advancing genome-wide association studies (GWAS) to identify common genetic factors that influence health and disease. For the purposes of this policy, a genome-wide association study is defined as any study of genetic variation across the entire human genome that is designed to identify genetic associations with observable traits (such as blood

U.S. Department of Health & Human Services www.hhs.gov
www.nih.gov

Genome-Wide Association Studies (GWAS)

Google™ Custom Search Search

GWAS Data Repository Policy

Introduction

Genome-wide association studies (GWAS) are used to identify common genetic factors that influence health and disease. In January 2008, the NIH implemented a policy for the sharing of data obtained in NIH-supported or conducted GWAS. The purpose of the policy is to foster science for the benefit of the public through the creation of a centralized NIH GWAS data repository. This Website supports the GWAS policy's implementation.

The NIH will continue to release additional information about the NIH GWAS policy on this site. Please e-mail questions about the policy to GWAS@mail.nih.gov.

In the Spotlight

NIH has implemented a new process for accessing certain aggregate datasets designated as general research use (GRU) through a single request for the [Compilation of Aggregate Genomic Data study](#). These data will continue to remain in controlled access with all the same expectations as other controlled-access data, including data management practices and annual reporting on data use.

For additional information about the new process and how to request access to the compilation of aggregate data, click [here](#).

Policy Announced: August 28, 2007
Effective Date: January 25, 2008

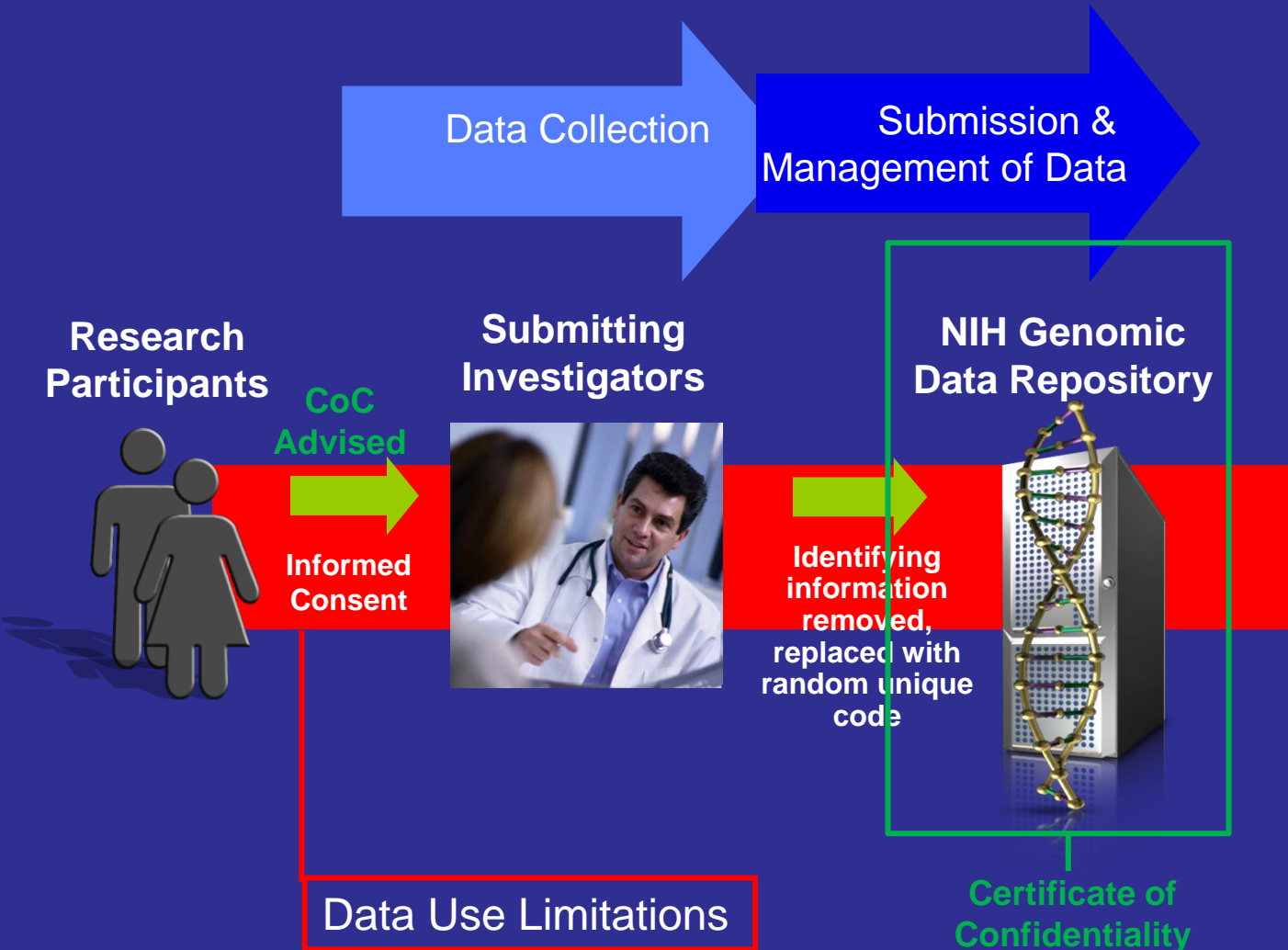
GWAS Homepage: <http://gwas.nih.gov/>

Guiding Principle for NIH Policies

The greatest public benefit will be realized if data from genomic studies are made available, under terms and conditions consistent with the informed consent provided by individual participants, in a timely manner to the largest possible number of investigators.

- Respect for Participants
- Data Sharing
- Freedom to Operate

NIH Data Management Overview

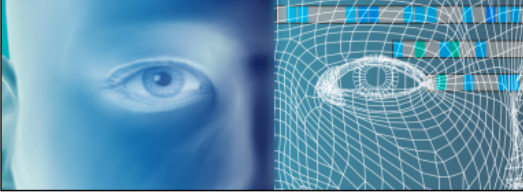


IT Security Measures

[NCBI](#) [Resources](#) [How To](#) My NCBI [Sign In](#)

dbGaP [Search](#)

[Limits](#) [Advanced](#) [Help](#)



dbGaP

The database of Genotypes and Phenotypes (dbGaP) was developed to archive and distribute the results of studies that have investigated the interaction of genotype and phenotype.

Getting Started

- [dbGaP Tutorial](#)
- [Overview](#)
- [FAQ](#)
- [How to Submit](#)
- [Browse Top Level Studies](#)

Access dbGaP Data

- [Apply for Controlled Access Data](#)
- [Public Data via ftp Download](#)
- [Association Results Browser](#)
- [Phenotype-Genotype Integrator](#)

Important Links

- [dbGaP RSS Feed](#)
- [Code of Conduct](#)
- [Security Procedures](#)
- [Contact Us](#)

Expectations for Data Users

Latest Studies

Study	Embargo Release	Details	Participants	Type Of Study	Links	Platform
phs000504.v1.p1 Genomic Sequencing of Medulloblastoma	Version 1:	V D A S	90	Case Set	Links	Agilent selected, 76bp paired end reads
phs000501.v1.p1 Compilation of Aggregate Genomic Data for General Research Use	Version 1:	V D A S	0	Aggregate Genomic Data		
phs000384.v1.p1 Genentech Whole Genome Sequencing of Four Hepatocellular Carcinoma Patients	Version 1: passed embargo	V D A S	4	Case Set	Links	Incorporated Assembler Version 1.8 HiSeq 2000
phs000500.v1.p1 Non-Invasive Whole Genome Sequencing of a Human Fetus	Version 1:	V D A S	0	Parent-Offspring Trios	Links	
phs000494.v1.p1 CCHMC - eMERGE Data	Version 1: 2013-07-23	V D A S	4562	Cohort	Links	AFFY_6.0 Human610_Quadv1_B Human660W-Quad_v1_A HumanOmni1_Quad_v1-0_B HumanOmni5-Quad

NIH Data Management Overview



**Research
Participants**



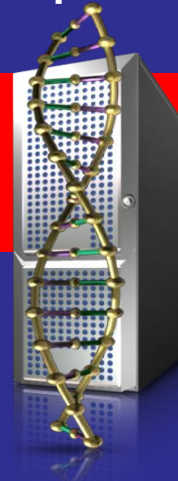
**Informed
Consent**

**Submitting
Investigators**



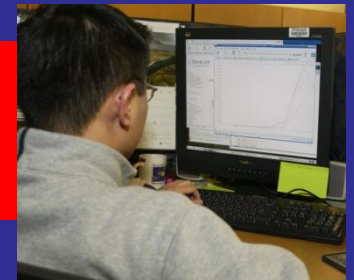
**Identifying
information
removed,
replaced with
random unique
code**

**NIH Genomic
Data Repository**



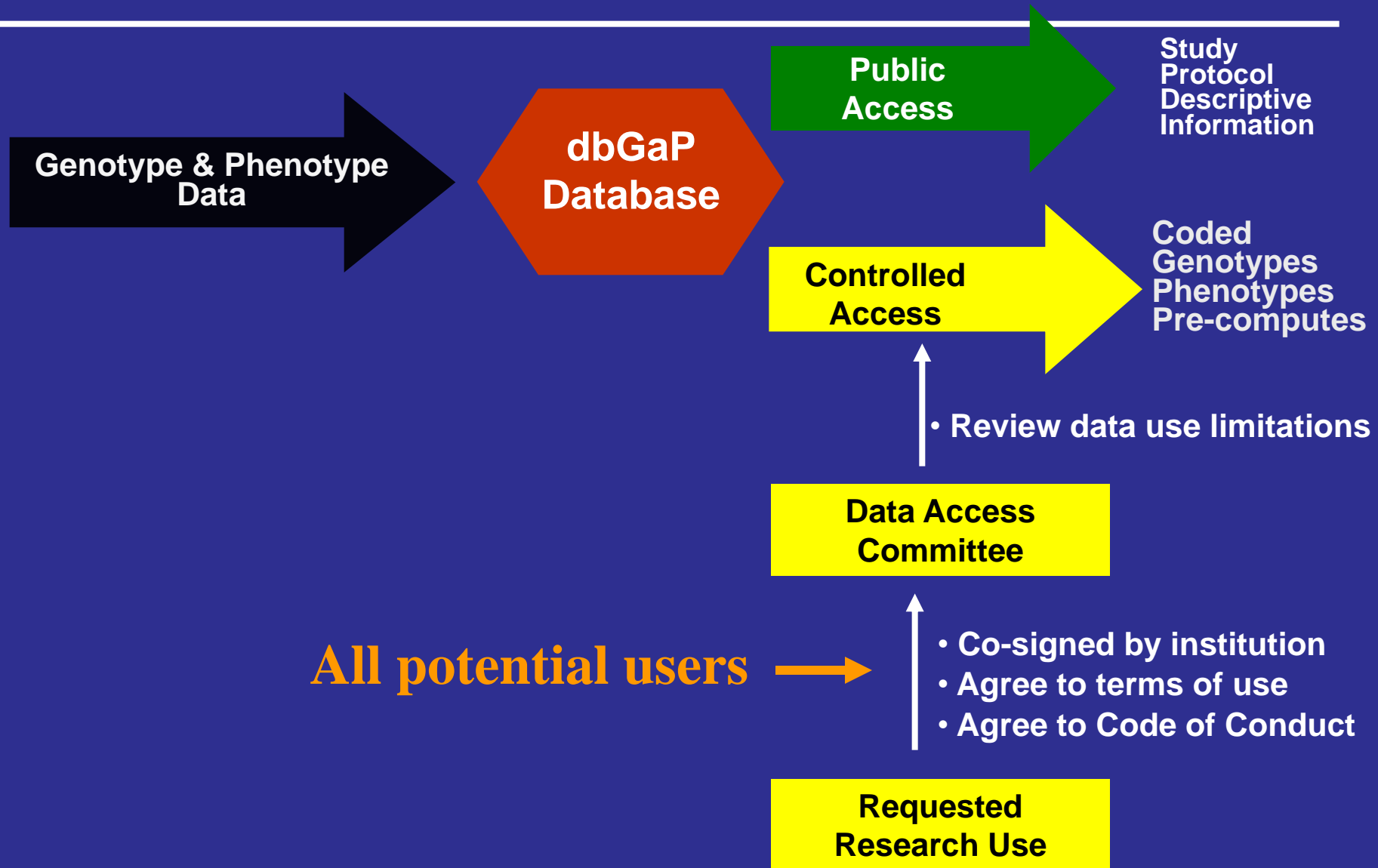
**Data
Access
Request by
Consent
Groups**

**Recipient
Investigators**



Data Use Limitations

Data Access



Data Use Certification Agreement

Terms and conditions include that requesters will:

- be responsible for compliance with federal, state, and local policies
- only use the data for the specified research use
- not identify study participants
- not transfer data beyond approved users
- immediately notify the DAC if a security breach occurs
- submit brief annual updates on research and publications
- be identified as an Approved User within the dbGaP
- acknowledge other GWAS policies
- abide by Genomic Data User Code of Conduct

Public Disclosure under FOIA

- dbGaP GWAS data will be coded and deidentified.
- Policy concern remains that the extensive genomic data in dbGaP is intrinsically unique.
- NIH intends to deny FOIA requests for individual-level genomic data.
- ACD Working Group recommended seeking a legislative exception for genomic data

Compelled Disclosure

- Law enforcement and other judicial requests for data are handled under different procedures than FOIA requests.
- Standard protection is provided through the Secretary's "301(d)" authority (*aka Certificates of Confidentiality*).
 - Certificates are issued for a single project.
 - Submitting Investigators are encouraged to consider requesting a Certificate of Confidentiality.
 - dbGaP has received a Certificate.

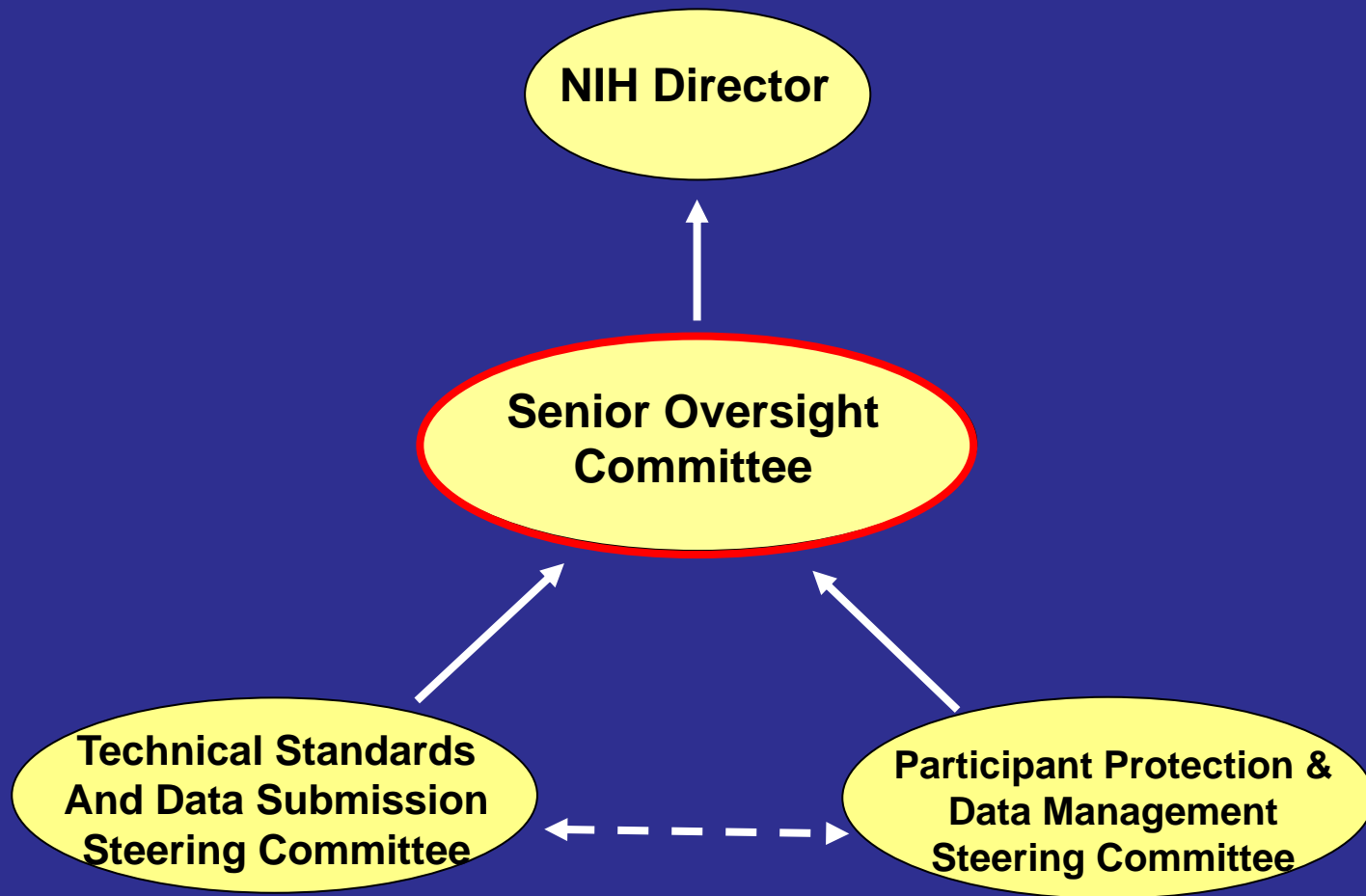
Exceptions to Data Deposition

- Policy notes that there will be cases where data deposition may not be appropriate
- Requests for exceptions are to come in within the grant Data Sharing Plan
- To date: 7 requests granted
 - Limited consent
 - Legal restrictions
 - Localized geographic representation

Data Use Experience

- Over 300 available studies (more in pipeline)
- Over 500 organizations across the research community and from 36 countries with Approved Users
- Number of active projects continues to increase substantially each year
- Estimates indicate >450 publications per year
- 3539 approved projects since dbGaP launch and there have been 16 Data Management Incidents

Stewardship & Oversight

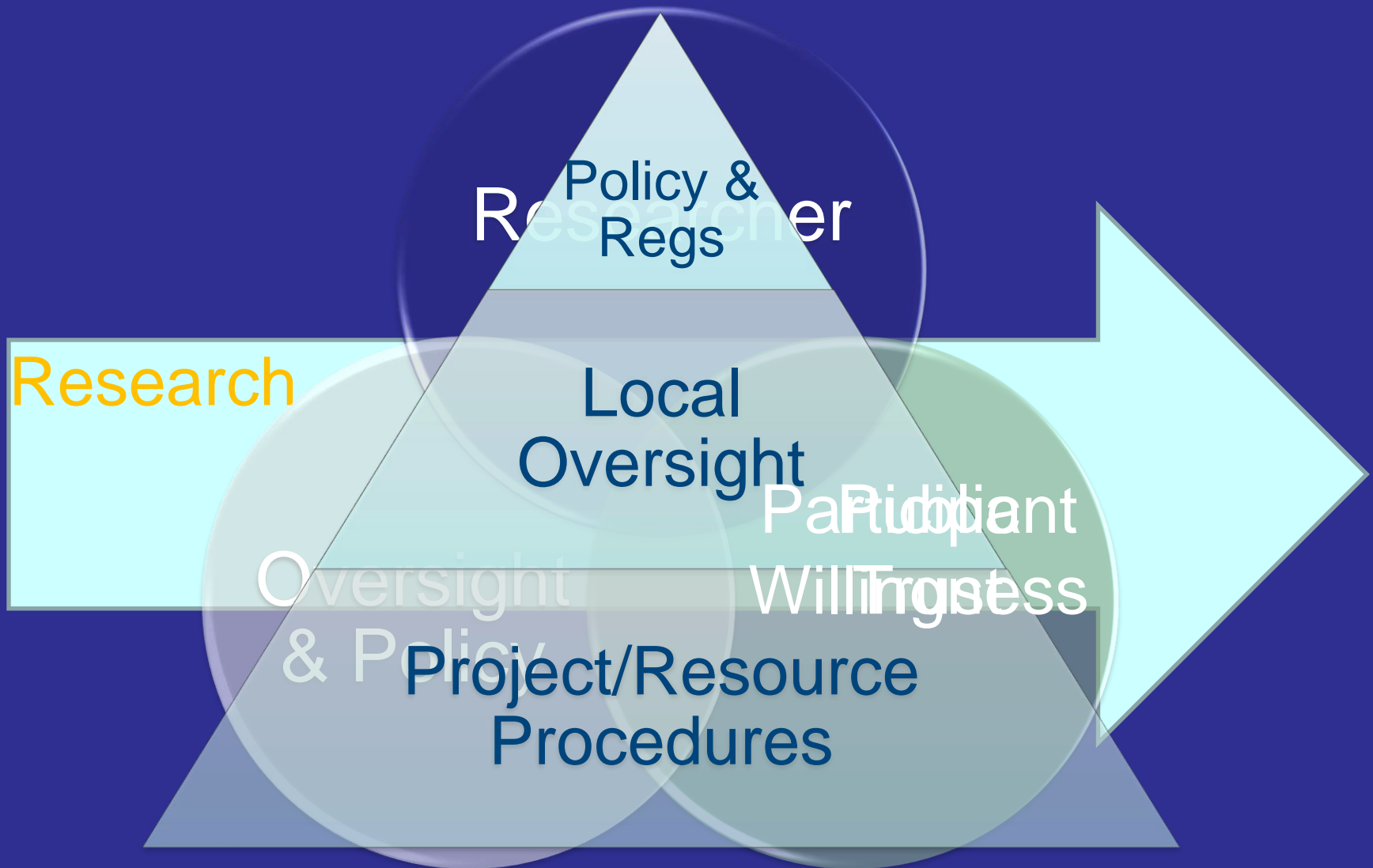


Fast Forward to 2012

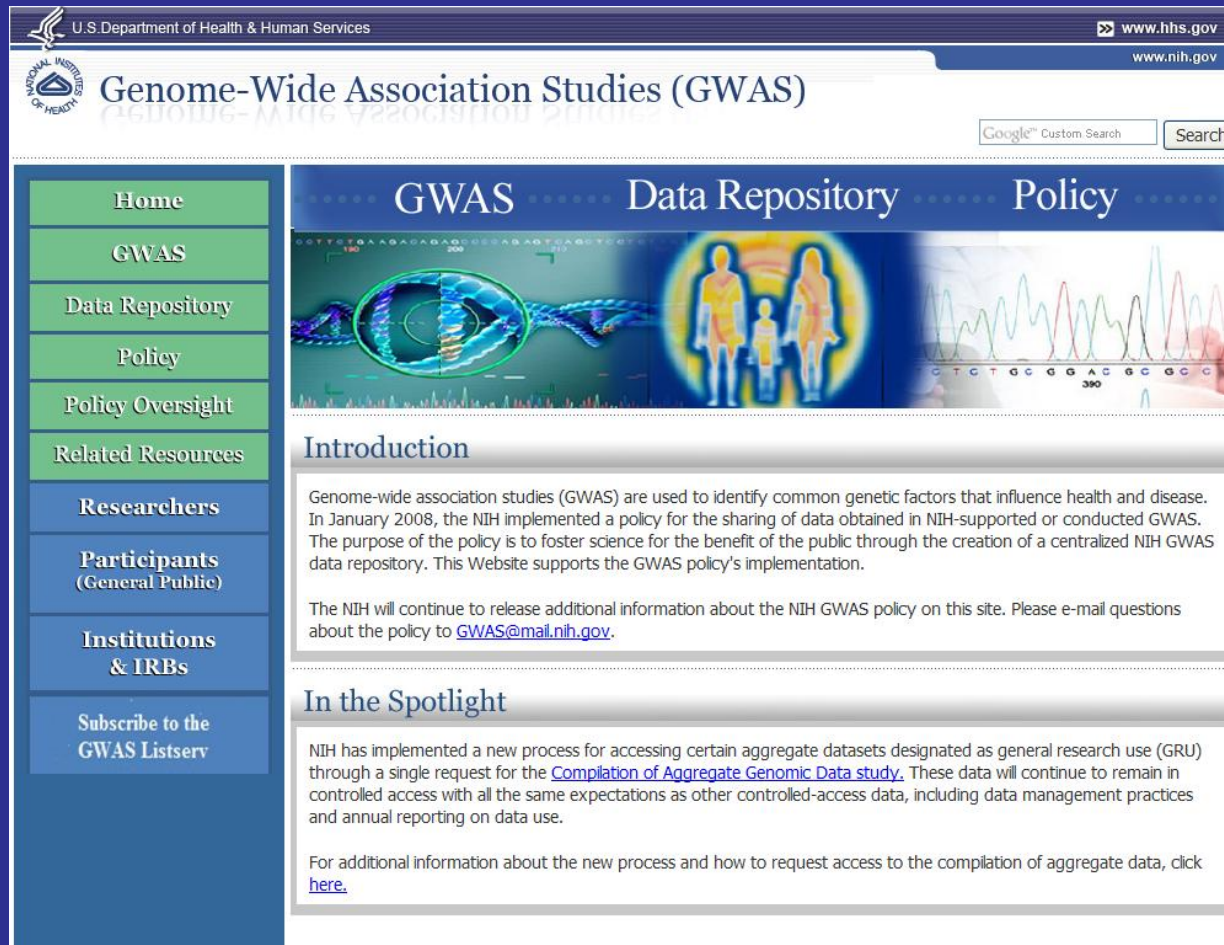


- The Power of Statistics
- Accessibility of Technology
- Participant-focused Research
- Participant Directions
- Media Coverage
- Policy Developments

Systems of Oversight



For more Information & Updates



The screenshot shows the official NIH website for Genome-Wide Association Studies (GWAS). The header includes the U.S. Department of Health & Human Services logo and the NIH logo, with the title "Genome-Wide Association Studies (GWAS)". Navigation links for "Home", "GWAS", "Data Repository", "Policy", "Policy Oversight", and "Related Resources" are on the left. The main content area features a banner with "GWAS Data Repository Policy" and an "Introduction" section. The introduction explains that GWAS are used to identify common genetic factors influencing health and disease, and that a centralized NIH GWAS data repository was created in January 2008. It also mentions that the NIH will continue to release additional information and that users can email questions to GWAS@mail.nih.gov. Below the introduction is an "In the Spotlight" section, which highlights a new process for accessing certain aggregate datasets designated as general research use (GRU) through a single request for the [Compilation of Aggregate Genomic Data study](#). It notes that these data will remain in controlled access with the same expectations as other controlled-access data, including data management practices and annual reporting on data use. A link [here](#) is provided for additional information about the new process and how to request access to the compilation of aggregate data.

U.S. Department of Health & Human Services www.hhs.gov
www.nih.gov

Genome-Wide Association Studies (GWAS)

Google Custom Search Search

Home
GWAS
Data Repository
Policy
Policy Oversight
Related Resources
Researchers
Participants (General Public)
Institutions & IRBs
Subscribe to the GWAS Listserv

GWAS Data Repository Policy

Introduction

Genome-wide association studies (GWAS) are used to identify common genetic factors that influence health and disease. In January 2008, the NIH implemented a policy for the sharing of data obtained in NIH-supported or conducted GWAS. The purpose of the policy is to foster science for the benefit of the public through the creation of a centralized NIH GWAS data repository. This Website supports the GWAS policy's implementation.

The NIH will continue to release additional information about the NIH GWAS policy on this site. Please e-mail questions about the policy to GWAS@mail.nih.gov.

In the Spotlight

NIH has implemented a new process for accessing certain aggregate datasets designated as general research use (GRU) through a single request for the [Compilation of Aggregate Genomic Data study](#). These data will continue to remain in controlled access with all the same expectations as other controlled-access data, including data management practices and annual reporting on data use.

For additional information about the new process and how to request access to the compilation of aggregate data, click [here](#).

<http://gwas.nih.gov>

email: laura.rodriguez@nih.gov or gwas@mail.nih.gov